



Widely Linear Equalization for IQ Imbalance and Skew Compensation in Optical Coherent Receivers

Porto da Silva, Edson; Zibar, Darko

Published in:
Journal of Lightwave Technology

Link to article, DOI:
[10.1109/JLT.2016.2577716](https://doi.org/10.1109/JLT.2016.2577716)

Publication date:
2016

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Porto da Silva, E., & Zibar, D. (2016). Widely Linear Equalization for IQ Imbalance and Skew Compensation in Optical Coherent Receivers. *Journal of Lightwave Technology*, 34(15), 3577-3586.
<https://doi.org/10.1109/JLT.2016.2577716>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Widely Linear Equalization for IQ Imbalance and Skew Compensation in Optical Coherent Receivers

Edson Porto da Silva, *Student Member, IEEE*, and Darko Zibar, *Member, IEEE*

Abstract—In this paper, an alternative approach to design linear equalization algorithms for optical coherent receivers is introduced. Using widely linear complex analysis, a general analytical model it is shown, where In-phase/quadrature (IQ) imbalances and IQ skew at the coherent receiver front-end are naturally included in channel equalization problem. Next, the problem of chromatic dispersion (CD) compensation after imbalanced and skewed coherent detection is analyzed. Based on the analytical models obtained it is demonstrated that, under the presence of such receiver front-end imperfections, the complexity of the channel equalization filter which is able to provide optimal performance, on the minimum mean square error (MMSE) sense, will scale proportionally to twice the complexity of CD compensation, if standard zero-forcing equalization of CD is applied as first equalization procedure. For the last, it is shown that, by applying the widely linear complex analysis, one can derive a complex-valued adaptive equalizer structure which is able to compensate for linear IQ-mixing effects at the receiver front-end. By extensive numerical simulations, the performance versus complexity of the proposed equalizer is shown to match the predictions of the analytical models derived.

Index Terms—Coherent receiver, Adaptive equalization

I. INTRODUCTION

Digital signal processing (DSP) has been extensively applied in coherent optical communications to enable equalization of linear fiber impairments, such as chromatic dispersion (CD) and polarization mode dispersion (PMD) [1]. The design of such equalization algorithms has been mostly based on the physical models of the fiber linear effects, and usually coherent receivers employ two steps of linear equalization: first the accumulated CD is compensated by a zero-forcing (ZF) static equalizer, and then a adaptive equalizer is used to compensate for residual CD and polarization mixing. Zero-forcing equalization is performed when the channel frequency response $H(\omega)$ is compensated by directly applying its inverse $H(\omega)^{-1} = 1/H(\omega)$ to the detected signals at the receiver [2]. The channel response is assumed to be known at the receiver. As a drawback, the use of such equalizers may result in performance degradation for systems with frequency-selective channel characteristics, due to noise enhancement. Fortunately, CD has all-pass frequency response characteristic and its compensation by a static ZF equalizer is practically viable for optical coherent receivers [1].

However, when such systems evolve to use high order modulation formats or high symbol rate signaling, imperfections

at the receiver front-end may have considerable influence on the transmission performance. Transceivers employing large order M -ary quadrature and amplitude modulation (QAM) formats usually require high signal-to-noise ratio (SNR) values for reliable transmission, which make their implementation less tolerant to receiver front-end imbalances. Therefore, two important problems in this scenario are how these imperfections relate to the performance of the signal detection and equalization, and how the DSP can be designed to mitigate eventual penalties.

Recent works have proposed adaptive equalization architectures robust to receiver front-end imperfections, such as In-phase/Quadrature (IQ) imbalance and IQ skew [3]–[6], here generally referred to as IQ-mixing effects. However, none of these references have presented a detailed analysis on the connection between the receiver IQ-mixing effects and the complexity requirements for equalization algorithms at the receiver. Mainly, the joint impact of ZF equalization of CD and IQ-mixing on the performance of the adaptive equalizer has not been completely clarified. For example, none of the references provide design rules specifying the required number of taps for the adaptive equalization algorithm.

As main contribution of this paper, the use of augmented complex-valued models, also known as widely linear (WL) complex-valued transformations, is proposed to define a general and systematic design for the receiver equalizers which takes into account receiver imperfections such as IQ-mixing effects. Following this approach, analytical models are derived to explain the interaction of ZF equalization of CD with receiver front-end imperfections. Based on such models the complexity requirements for the minimum mean square error (MMSE) equalization algorithms are extracted. It is shown that the general solution for the equalization problem correspond to a WL complex-valued adaptive equalizer structure, which is proposed and validated with extensive numerical simulations.

Complex-valued WL transformations have been extensively investigated in the DSP literature, particularly for statistical modeling and equalization problems. A comprehensive overview on this topic can be found in reference [9]. Examples of application of complex-valued WL equalization techniques can be found in the following references [10]–[13].

This paper is divided as follows: in section II, a formal description of the connection between linear real-valued and widely linear complex-valued transformations is presented. In section III, it is shown how IQ-mixing effects can be seen as complex-valued WL transformations. Next, it is shown that this approach lead to closed expressions that describe the resultant effect of IQ-mixing when the receiver performs ZF equalization of CD. Based on the obtained expressions,

The authors are with the Department of Photonics Engineering (DTU Fotonik), Technical University of Denmark - DTU, Kgs. Lyngby, Denmark, e-mails: {edpod, dazi}@fotonik.dtu.dk.

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

the complexity requirements are discussed for the adaptive equalizer to be able to recover the signals with no penalties coming from skew or imbalance at the receiver front-end. In section IV, the general structure of a WL complex-valued adaptive equalizer for a polarization multiplexed receiver is derived. In section V, extensive simulation results and discussions are presented, comparing the performance of the proposed equalizer with a standard linear complex-valued equalizer, which lead to the conclusions.

II. AUGMENTED COMPLEX-VALUED SIGNAL PROCESSING

We start with definitions which will be necessary to formally establish equivalence relations between linear transformations in real vector spaces and complex vector spaces. Here we use the same approach of [9]. Let $\mathbf{u}_r \in \mathbb{R}^{N \times 1}$ and $\mathbf{u}_i \in \mathbb{R}^{N \times 1}$. With \mathbf{u}_r and \mathbf{u}_i we can construct three related vectors:

- I. The real composite vector: $\mathbf{u}_{\mathbb{R}} = [\mathbf{u}_r^T, \mathbf{u}_i^T]^T = \begin{bmatrix} \mathbf{u}_r \\ \mathbf{u}_i \end{bmatrix}$.
- II. The complex vector: $\mathbf{u} = \mathbf{u}_r + j\mathbf{u}_i$.
- III. The augmented complex vector: $\underline{\mathbf{u}} = [\mathbf{u}^T, \mathbf{u}^H]^T = \begin{bmatrix} \mathbf{u} \\ \mathbf{u}^* \end{bmatrix}$.

where $\mathbf{u}_{\mathbb{R}} \in \mathbb{R}^{2N \times 1}$, $\mathbf{u} \in \mathbb{C}^{N \times 1}$ and $\underline{\mathbf{u}} \in \mathbb{C}_*^{2N \times 1}$. The conjugate operation is denoted by $(\cdot)^*$, the transpose operation is denoted by $(\cdot)^T$, and the conjugate transpose (hermitian) is denoted by $(\cdot)^H$. The notation $\mathbb{C}_*^{2N \times 1}$ refers to the augmented version of the set $\mathbb{C}^{2N \times 1}$.

The mapping between the complex augmented vector $\underline{\mathbf{u}}$ and the real composite vector $\mathbf{u}_{\mathbb{R}}$ is given by the linear real-to-complex transformation

$$\mathbf{T}_N = \begin{bmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{bmatrix} \in \mathbb{C}^{2N \times 2N} \implies \underline{\mathbf{u}} = \mathbf{T}_N \mathbf{u}_{\mathbb{R}} \quad (1)$$

where \mathbf{T}_N has the property $\mathbf{T}_N \mathbf{T}_N^H = \mathbf{T}_N^H \mathbf{T}_N = 2\mathbf{I}$, i.e. $\frac{1}{\sqrt{2}}\mathbf{T}_N$ is a unitary matrix.

Now suppose we apply an arbitrary linear transformation $\mathbf{M} : \mathbb{R}^{2N \times 1} \rightarrow \mathbb{R}^{2N \times 1}$ to the vector $\mathbf{u}_{\mathbb{R}}$

$$\mathbf{v}_{\mathbb{R}} = \mathbf{M} \mathbf{u}_{\mathbb{R}} \rightarrow \begin{bmatrix} \mathbf{v}_r \\ \mathbf{v}_i \end{bmatrix} = \begin{bmatrix} \mathbf{m}_{11} & \mathbf{m}_{12} \\ \mathbf{m}_{21} & \mathbf{m}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{u}_r \\ \mathbf{u}_i \end{bmatrix} \quad (2)$$

where $\mathbf{m}_{ij} \in \mathbb{R}^{N \times N}$. The transformation \mathbf{M} is equivalent to a complex augmented transformation $\underline{\mathbf{H}} : \mathbb{C}_*^{2N \times 1} \rightarrow \mathbb{C}_*^{2N \times 1}$ which can be derived as follows

$$\begin{aligned} \mathbf{v}_{\mathbb{R}} &= \mathbf{M} \mathbf{u}_{\mathbb{R}} \Rightarrow \mathbf{T}_N \mathbf{v}_{\mathbb{R}} = \mathbf{T}_N \mathbf{M} \frac{1}{2} \mathbf{T}_N^H \mathbf{T}_N \mathbf{u}_{\mathbb{R}} \\ \underline{\mathbf{v}} &= \frac{1}{2} \mathbf{T}_N \mathbf{M} \mathbf{T}_N^H \underline{\mathbf{u}} \Rightarrow \underline{\mathbf{v}} = \underline{\mathbf{H}} \underline{\mathbf{u}} \\ \underline{\mathbf{H}} &= \frac{1}{2} \mathbf{T}_N \mathbf{M} \mathbf{T}_N^H \end{aligned} \quad (3)$$

$$\quad (4)$$

where the matrix $\underline{\mathbf{H}} \in \mathbb{C}^{2N \times 2N}$ has the particular block pattern

$$\underline{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{H}_2^* & \mathbf{H}_1^* \end{bmatrix} \quad (5)$$

with

$$\mathbf{H}_1 = \frac{1}{2}[\mathbf{m}_{11} + \mathbf{m}_{22} + j(\mathbf{m}_{21} - \mathbf{m}_{12})] \quad (6)$$

$$\mathbf{H}_2 = \frac{1}{2}[\mathbf{m}_{11} - \mathbf{m}_{22} + j(\mathbf{m}_{21} + \mathbf{m}_{12})] \quad (7)$$

From $\underline{\mathbf{H}}$ and the definition of the augmented complex vectors $\underline{\mathbf{v}}$ and $\underline{\mathbf{u}}$, we can derive the general form of a *widely linear* transformation in \mathbb{C}^N

$$\mathbf{v} = \mathbf{H}_1 \mathbf{u} + \mathbf{H}_2 \mathbf{u}^*. \quad (8)$$

We denote by *strictly linear* transformation in \mathbb{C}^N the particular case of Eq. (8) when $\mathbf{H}_2 = 0$. Therefore, one can directly conclude that a linear transformation in \mathbb{R}^{2N} can only be completely described by a strictly linear transformation in \mathbb{C}^N if the condition $\mathbf{H}_2 = 0$ is satisfied. From Eq. (7), this condition implies that \mathbf{M} in Eq. (2) has to have the form

$$\mathbf{M} = \begin{bmatrix} \mathbf{m}_{11} & \mathbf{m}_{12} \\ -\mathbf{m}_{12} & \mathbf{m}_{11} \end{bmatrix} \quad (9)$$

Most of the conventional complex-valued DSP implementations of equalizers (static and adaptive) for optical coherent receivers adopt strictly linear structures, since passive linear propagation effects happening into the fiber are completely described by strictly linear complex-valued models [7]. However, as it is shown in the next sections, Eq. (8) has implications on the equalization performance if additional effects that violate Eq. (9) are present between transmitter and receiver. In particular, the focus will be directed to equalization problem when the receiver front-end is subject to IQ imbalance or IQ skew.

III. IMBALANCED RECEIVER MODELS

In this section, complex-valued analytical models of IQ imbalance and IQ skew effects at the receiver are derived. The models obtained apply to a single polarization receiver. Nevertheless, considering that CD and PMD are strictly linear in the complex domain (i.e. when described in the complex domain, they satisfy the condition $\mathbf{H}_2 = 0$), such models can be directly extended for a polarization multiplexed receiver. Although this fact will not be proved here, it can be easily verified following the analysis presented in reference [7]. No imbalance or skew is considered at the transmitter side. Frequency offset between transmitter and receiver is not included in the model derivations, although it is included in the simulation setup and results. Here the notations $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ are used to denote the direct and the inverse Fourier transform operation, respectively. The notation of matrix multiplication, for the sake of clarity, is also extended to denote the convolution operation, as it is commonly used in the modeling of multiple-input multiple-output (MIMO) systems. The meaning should be clear by the context.

A. Receiver IQ Imbalance Model

Here it is assumed the case where the IQ imbalance is present only at the receiver side. We can model the IQ imbalance writing the equivalent complex baseband of an ideal local oscillator (LO) laser as

$$\begin{aligned} m_{lo}^{rx}(t) &= (1 - \epsilon_{rx}) \cos(\omega_c t - \phi_{rx}/2) \\ &\quad - j(1 + \epsilon_{rx}) \sin(\omega_c t + \phi_{rx}/2) \\ &= k_1 e^{-j\omega_c t} + k_2 e^{j\omega_c t} \end{aligned} \quad (10)$$

with

$$\begin{aligned} k_1 &= [(1 - \epsilon_{rx})e^{j\frac{\phi_{rx}}{2}} + (1 + \epsilon_{rx})e^{-j\frac{\phi_{rx}}{2}}]/2 \\ &= k_{1,r} + jk_{1,i} \end{aligned} \quad (11)$$

$$\begin{aligned} k_2 &= [(1 - \epsilon_{rx})e^{-j\frac{\phi_{rx}}{2}} - (1 + \epsilon_{rx})e^{j\frac{\phi_{rx}}{2}}]/2 \\ &= k_{2,r} + jk_{2,i} \end{aligned} \quad (12)$$

where $\omega_c = 2\pi f_c$, and f_c is the carrier frequency. The parameters $\{\epsilon_{rx}, \phi_{rx}\}$ represent the effective amplitude and phase imbalance of the receiver, respectively [8]. Let $s(t) = s_I(t) + js_Q(t)$ be the equivalent complex baseband of the received signal at the instant t , then

$$\hat{s}(t) = k_1 s(t) + k_2 s^*(t) \quad (13)$$

where $\hat{s}(t)$ is the resultant signal obtained after imbalanced detection of $s(t)$. Equation (13) can be rewritten in the real-valued matrix form

$$\begin{aligned} \begin{bmatrix} \hat{s}_I(t) \\ \hat{s}_Q(t) \end{bmatrix} &= \begin{bmatrix} k_{1,r} + k_{2,r} & -(k_{1,i} - k_{2,i}) \\ k_{1,i} + k_{2,i} & k_{1,r} - k_{2,r} \end{bmatrix} \begin{bmatrix} s_I(t) \\ s_Q(t) \end{bmatrix} \Rightarrow \\ \begin{bmatrix} \hat{s}_I(t) \\ \hat{s}_Q(t) \end{bmatrix} &= \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} s_I(t) \\ s_Q(t) \end{bmatrix} \end{aligned} \quad (14)$$

Equation (14) is a linear transformation included in the set defined in Eq. (2). However, in the general case, it does not comply with the form of Eq. (9). Hence, we conclude that IQ imbalance can only be completely described in the complex domain by a widely linear transformation.

B. Receiver IQ Time Skew Model

Assuming that the IQ components of the received signal $s(t)$ are relatively skewed by a delay τ , the resultant signal $\hat{s}(t)$ can be described by

$$\hat{s}(t) = s_I(t) * \delta(t - \tau/2) + js_Q(t) * \delta(t + \tau/2) \quad (15)$$

which can be rewritten in the real-valued matrix form

$$\begin{bmatrix} \hat{s}_I(t) \\ \hat{s}_Q(t) \end{bmatrix} = \begin{bmatrix} \delta(t - \tau/2) & 0 \\ 0 & \delta(t + \tau/2) \end{bmatrix} \begin{bmatrix} s_I(t) \\ s_Q(t) \end{bmatrix}. \quad (16)$$

Also Eq. (16) belongs to the set of linear transformations defined in Eq. (2) but it is not strictly linear, as it does not comply with Eq. (9). Therefore, IQ skew is also a widely linear transformation in the complex domain, which can be represented by

$$\hat{s}(t) = d_1(t) * s(t) + d_2(t) * s^*(t) \quad (17)$$

where

$$\begin{aligned} d_1(t) &= [\delta(t - \tau/2) + \delta(t + \tau/2)]/2 \\ d_2(t) &= [\delta(t - \tau/2) - \delta(t + \tau/2)]/2. \end{aligned}$$

C. Chromatic Dispersion Compensation with IQ Imbalance at the Receiver

Here we assume that the signal $s(t)$ has propagated over a distance L in a single mode fiber, and that the receiver attempts to equalize the CD immediately after the coherent detection. The frequency response of the CD is given by

$$H_{cd}(\omega) = e^{-j\frac{\beta_2(\omega_c)}{2}\omega^2 L}. \quad (18)$$

Let $h_{cd}(t) = \mathcal{F}^{-1}\{H_{cd}(\omega)\}$ be the complex-valued impulse response of the chromatic dispersion, we can write

$$h_{cd}(t) = h_{cd,I}(t) + jh_{cd,Q}(t). \quad (19)$$

and $\hat{s}(t) = h_{cd}(t) * s(t)$, which corresponds to

$$\begin{bmatrix} \hat{s}_I(t) \\ \hat{s}_Q(t) \end{bmatrix} = \begin{bmatrix} h_{cd,I}(t) & -h_{cd,Q}(t) \\ h_{cd,Q}(t) & h_{cd,I}(t) \end{bmatrix} \begin{bmatrix} s_I(t) \\ s_Q(t) \end{bmatrix} \quad (20)$$

If IQ imbalance is present at the receiver, we can use Eq. (14) and Eq. (20) to obtain

$$\begin{bmatrix} \hat{s}_I(t) \\ \hat{s}_Q(t) \end{bmatrix} = \begin{bmatrix} k_{11}h_{cd,I}(t) + k_{12}h_{cd,Q}(t) \\ k_{21}h_{cd,I}(t) + k_{22}h_{cd,Q}(t) \\ -k_{11}h_{cd,Q}(t) + k_{12}h_{cd,I}(t) \\ -k_{21}h_{cd,Q}(t) + k_{22}h_{cd,I}(t) \end{bmatrix} \begin{bmatrix} s_I(t) \\ s_Q(t) \end{bmatrix} \quad (21)$$

From Eq. (21), the equivalent signal $\hat{s}(t)$ which will be forwarded to the CD equalizer is given by

$$\hat{s}(t) = m_1(t) * s(t) + m_2(t) * s^*(t) \quad (22)$$

where

$$\begin{aligned} m_1(t) &= [k_{11} + k_{22} + j(k_{21} - k_{12})]h_{cd}(t)/2 \\ &= K_1 h_{cd}(t) \\ m_2(t) &= [k_{11} - k_{22} + j(k_{21} + k_{12})]h_{cd}^*(t)/2 \\ &= K_2 h_{cd}^*(t). \end{aligned}$$

Since CD compensation is usually done in the frequency domain, it is convenient to write the equivalent system of Eq. (22) as

$$\hat{S}(\omega) = M_1(\omega)S(\omega) + M_2(\omega)S^*(-\omega) \quad (23)$$

where $S(\omega) = \mathcal{F}\{s(t)\}$, $S^*(-\omega) = \mathcal{F}\{s^*(t)\}$, and

$$\begin{aligned} M_1(\omega) &= K_1 H_{cd}(\omega) \\ M_2(\omega) &= K_2 H_{cd}^*(-\omega). \end{aligned}$$

Finally, assuming that the receiver applies the standard ZF equalizer $H_{cd}^{-1}(\omega) = H_{cd}^*(\omega) = H_{cd}^*(-\omega)$ to compensate the CD, it results in

$$\begin{aligned} \hat{S}_{ZF}(\omega) &= H_{cd}^{-1}(\omega)\hat{S}(\omega) \\ &= G_1(\omega)S(\omega) + G_2(\omega)S^*(-\omega) \end{aligned} \quad (24)$$

where

$$\begin{aligned} G_1(\omega) &= K_1 \\ G_2(\omega) &= K_2 [H_{cd}^*(\omega)]^2. \end{aligned}$$

D. Chromatic Dispersion Compensation with IQ Skew at the Receiver

Considering that the receiver adds IQ time skew to a dispersed signal, from Eq. (16) and Eq. (20), and following the same steps of subsection C, we obtain

$$\begin{bmatrix} \hat{s}_I(t) \\ \hat{s}_Q(t) \end{bmatrix} = \begin{bmatrix} h_{cd,I}(t - \tau/2) & -h_{cd,Q}(t - \tau/2) \\ h_{cd,Q}(t + \tau/2) & h_{cd,I}(t + \tau/2) \end{bmatrix} \begin{bmatrix} s_I(t) \\ s_Q(t) \end{bmatrix} \quad (25)$$

Therefore, $\hat{s}(t)$ can be expressed as

$$\hat{s}(t) = h_1(t) * s(t) + h_2(t) * s^*(t) \quad (26)$$

where

$$\begin{aligned} h_1(t) &= \{h_{cd,I}(t - \tau/2) + h_{cd,I}(t + \tau/2) \\ &\quad + j[h_{cd,Q}(t + \tau/2) + h_{cd,Q}(t - \tau/2)]\}/2 \\ h_2(t) &= \{h_{cd,I}(t - \tau/2) - h_{cd,I}(t + \tau/2) \\ &\quad + j[h_{cd,Q}(t + \tau/2) - h_{cd,Q}(t - \tau/2)]\}/2. \end{aligned}$$

The equivalent system of Eq. (26) in the frequency domain can be written as

$$\hat{S}(\omega) = H_1(\omega)S(\omega) + H_2(\omega)S^*(-\omega) \quad (27)$$

where

$$\begin{aligned} H_1(\omega) &= \{H_{cd,I}(\omega)e^{-j\omega\tau/2} + H_{cd,I}(\omega)e^{j\omega\tau/2} \\ &\quad + j[H_{cd,Q}(\omega)e^{j\omega\tau/2} + H_{cd,Q}(\omega)e^{-j\omega\tau/2}]\}/2 \\ &= H_{cd}(\omega) \cos(\omega\tau/2) \\ H_2(\omega) &= \{H_{cd,I}(\omega)e^{-j\omega\tau/2} - H_{cd,I}(\omega)e^{j\omega\tau/2} \\ &\quad + j[H_{cd,Q}(\omega)e^{j\omega\tau/2} - H_{cd,Q}(\omega)e^{-j\omega\tau/2}]\}/2 \\ &= -jH_{cd}^*(-\omega) \sin(\omega\tau/2). \end{aligned}$$

Again, assuming ZF compensation of CD, the result is

$$\begin{aligned} \hat{S}_{ZF}(\omega) &= H_{cd}^{-1}(\omega)\hat{S}(\omega) \\ &= G_1(\omega)S(\omega) + G_2(\omega)S^*(-\omega) \end{aligned} \quad (28)$$

where

$$\begin{aligned} G_1(\omega) &= \cos(\omega\tau/2) \\ G_2(\omega) &= -j[H_{cd}^*(\omega)]^2 \sin(\omega\tau/2). \end{aligned}$$

E. Discussion on the Implications of the Models

Here we assume that after ZF equalization of CD, the signal is forwarded to an adaptive equalizer which is responsible to demultiplex polarizations and to approximate the matched filter. Analyzing the expressions obtained in Eq. (24) and Eq. (28) we can point out the following remarks:

1. Since $G_1(\omega)$ and $G_2(\omega)$ are both strictly linear complex-valued transfer functions, the signal obtained after ZF compensation of CD is only completely described by a widely linear complex operation on $S(\omega)$. By observing this fact, we can infer that a standard strictly linear complex-valued equalizer will present a poor performance in recovering $S(\omega)$, since its structure can not completely represent the widely linear

transformation operated on the signal. Therefore, optimal equalization performance should be achieved only by widely linear complex-valued equalizer structure, or by its correspondence in the real-valued two dimensional space.

2. In fact, the impact of the conjugate term on the equalizer performance can even be quantified if we consider that it should be proportional to the signal to interference power ratio, where the power of the interference can be calculated from the term $G_2(\omega)S^*(-\omega)$ [8].
3. To optimally recover $S(\omega)$, the adaptive equalizer must be able to jointly process $\hat{S}_{ZF}(\omega)$ and $\hat{S}_{ZF}^*(-\omega)$, in order to invert the transfer functions $G_1(\omega)$ and $G_2(\omega)$. However, the impulse response $g_2(t) = \mathcal{F}^{-1}\{G_2(\omega)\}$ is twice as long as the response of the chromatic dispersion. Therefore, the complexity required to equalize such signal will scale proportionally to twice the complexity of zero-forcing CD compensation, independently if the equalizer is implemented with all real-valued or all complex-valued filters. This interesting conclusion explains why in [5] and [6] four (instead of the usual two) independent CD compensation complex-valued filters are necessary to enable a low complexity adaptive equalizer to compensate the IQ skew at the receiver, and why in [4] a large number of taps is required in order to recover the signal with the adaptive equalizer. The equalizer number of taps is also the limitation of [3], which fails when the signal has a large value of accumulated CD, as indicated in [6]. In fact, we can point out the equalizer structures proposed in [3]–[6] are equivalent in the sense that, giving the necessary number of equalizer taps to solve the equalization problem, they should provide similar performance.

F. On the WL-ZF equalization of IQ-mixing and CD

The presence of the conjugate terms in Eq. (23) and (27) implies that a complex-valued strictly linear ZF equalizer will not be able to completely eliminate the effect of the CD, since it cannot simultaneously compensate for $H_{cd}(\omega)$ and $H_{cd}^*(\omega)$. Another approach for this problem would be to design a WL-ZF equalizer of CD which includes the effects of IQ-mixing, such equalizer structure can be found by inverting the matrices in Eq. (21) and (25). In general, for a WL system described by

$$\hat{S}(\omega) = C_1(\omega)S(\omega) + C_2(\omega)S^*(-\omega) \quad (29)$$

assuming that $C_1(\omega)$ and $C_2(\omega)$ are invertible, the ZF widely linear equalizer will be given by

$$\begin{aligned} \hat{S}_{WL-ZF}(\omega) &= \frac{C_1^*(-\omega)}{C_1(\omega)C_1^*(-\omega) - C_2(\omega)C_2^*(-\omega)} \hat{S}(\omega) \\ &\quad - \frac{C_2(\omega)}{C_1(\omega)C_1^*(-\omega) - C_2(\omega)C_2^*(-\omega)} \hat{S}^*(-\omega). \end{aligned} \quad (30)$$

Note that if $C_2(\omega) = 0$, then

$$\hat{S}_{WL-ZF}(\omega) = \frac{1}{C_1(\omega)} \hat{S}(\omega) \quad (31)$$

which corresponds to the strictly linear ZF equalizer of $C_1(\omega)$. A necessary condition for the system in Eq. (30) to be invertible is $D(\omega) = C_1(\omega)C_1^*(-\omega) - C_2(\omega)C_2^*(-\omega) \neq 0$ for all ω . It is easy to check that for the systems in Eq. (27) and Eq. (28) $D(\omega) = 1$ always, which shows that any combination of IQ skew and CD can be fully compensated. Of course, since the design of the WL-ZF equalizer requires the receiver to know exactly the IQ imbalance of IQ skew at the front-end, the best procedure would be to compensate the IQ-mixing before the equalization blocks and save DSP complexity, instead. Nevertheless, Eq. (30) is useful to comprehend what an complex-valued WL adaptive equalizer should look like to recover a signal impaired by IQ-mixing and CD.

A hybrid approach to solve the equalization problem is also possible, as indirectly demonstrated in references [5] and [6]. There the authors propose to modify the ZF equalization of CD and to take advantage of symmetries of the CD impulse response to reduce complexity, further using a 4×2 complex-valued adaptive equalizer to compensate for the IQ skew. Their approach can be seen as the result of an direct application of Eq. (30) to find the WL equalizer for the system in Eq. (27). However, this approach may impose challenges to the synchronization and timing recovery algorithms, since they have to be performed before compensation of CD.

IV. ADAPTIVE EQUALIZER USING AUGMENTED JONES VECTORS

Optical coherent receivers commonly employ finite impulse response (FIR) MIMO adaptive equalizers to perform polarization demultiplexing and to approximate the matched filtered detection. In the previous sections it was shown how the augmented complex-valued signal processing can be used to obtain models in the complex domain which properly represent IQ-mixing related effects. The aim of this section is to derive a general complex-valued MIMO adaptive equalizer structure which should be able to compensate for reversible IQ-mixing effects. From now on the signals are assumed to be discrete time versions of their respective continuous time waveforms, sampled at least to the double of the baud rate. The elements in the matrix representation of the equalizers are FIR filters.

A. Equalizer General Structure

The structure of the most general adaptive equalizer for an polarization multiplexed receiver is given by

$$\begin{bmatrix} \hat{x}_I \\ \hat{x}_Q \\ \hat{y}_I \\ \hat{y}_Q \end{bmatrix} = \begin{bmatrix} h_{x_I x_I}^T & h_{x_I x_Q}^T & h_{x_I y_I}^T & h_{x_I y_Q}^T \\ h_{x_Q x_I}^T & h_{x_Q x_Q}^T & h_{x_Q y_I}^T & h_{x_Q y_Q}^T \\ h_{y_I x_I}^T & h_{y_I x_Q}^T & h_{y_I y_I}^T & h_{y_I y_Q}^T \\ h_{y_Q x_I}^T & h_{y_Q x_Q}^T & h_{y_Q y_I}^T & h_{y_Q y_Q}^T \end{bmatrix} \begin{bmatrix} x_I \\ x_Q \\ y_I \\ y_Q \end{bmatrix} \quad (32)$$

It corresponds to a 4×4 real-valued MIMO system where the inputs are the real and imaginary parts of the received complex-valued polarization signals (x, y) : $x_I = \text{Re}(x)$,

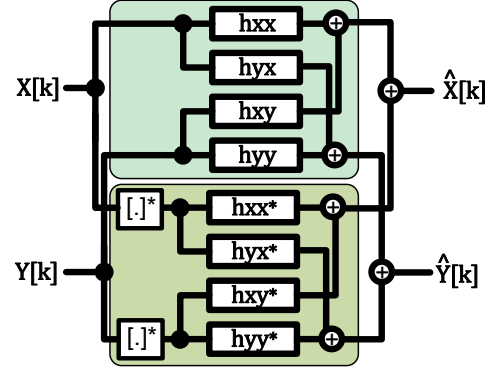


Fig. 1 Block diagram of the 2×2 MIMO complex-valued WL adaptive equalizer.

$x_Q = \text{Im}(x)$, $y_I = \text{Re}(y)$, $y_Q = \text{Im}(y)$. For convenience, define

$$\mathbf{M}_4 = \begin{bmatrix} h_{x_I x_I}^T & h_{x_I x_Q}^T & h_{x_I y_I}^T & h_{x_I y_Q}^T \\ h_{x_Q x_I}^T & h_{x_Q x_Q}^T & h_{x_Q y_I}^T & h_{x_Q y_Q}^T \\ h_{y_I x_I}^T & h_{y_I x_Q}^T & h_{y_I y_I}^T & h_{y_I y_Q}^T \\ h_{y_Q x_I}^T & h_{y_Q x_Q}^T & h_{y_Q y_I}^T & h_{y_Q y_Q}^T \end{bmatrix} \quad (33)$$

The equalizer structure in Eq. (32) should be able to compensate a general reversible channel response that can be entirely represented as a linear transformation in \mathbb{R}^4 . It is possible to show that, in absence of IQ-mixing, such equalizer can be reduced to the standard 2×2 complex-valued equalizer defined by

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} h_{xx}^H & h_{xy}^H \\ h_{yx}^H & h_{yy}^H \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (34)$$

where $\mathbf{s} = [x \ y]^T$ and $\hat{\mathbf{s}} = [\hat{x} \ \hat{y}]^T$ are the Jones vector representation of equalizer's input and output, respectively. Note that $\hat{\mathbf{s}} \in \mathbb{C}^{2 \times 1}$, while $\mathbf{s} \in \mathbb{C}^{2N \times 1}$ and the \mathbf{h} -elements of $\mathbf{M}_4 \in \mathbb{R}^{N \times 1}$, where N is the number of taps of the equalizer's FIR structure.

Equation (34) follows directly from the strict linearity of the CD and PMD in the complex domain. However, in the general case, the structures of Eq. (32) and Eq. (34) are not equivalent. In order to derive the equivalent version of Eq. (32) in the complex domain, we perform the linear transformation defined in Eq. (4). In this case, the \mathbf{T}_N transformation defined in Eq. (1) resolves to

$$\mathbf{T}_2 = \begin{bmatrix} 1 & j & 0 & 0 \\ 1 & -j & 0 & 0 \\ 0 & 0 & 1 & j \\ 0 & 0 & 1 & -j \end{bmatrix}. \quad (35)$$

Note that we have linearly rearranged the columns and rows of \mathbf{T}_2 in order to clarify the following steps, and that this linear transformation does not change its unitary properties. \mathbf{T}_2 can then be used to define the *augmented* Jones vector \mathbf{s} , which is obtained by

$$\mathbf{s} = \begin{bmatrix} x \\ x^* \\ y \\ y^* \end{bmatrix} = \begin{bmatrix} 1 & j & 0 & 0 \\ 1 & -j & 0 & 0 \\ 0 & 0 & 1 & j \\ 0 & 0 & 1 & -j \end{bmatrix} \begin{bmatrix} x_I \\ x_Q \\ y_I \\ y_Q \end{bmatrix} \quad (36)$$

Directly applying Eq. (4) to \mathbf{T}_2 and \mathbf{M}_4 , and after some algebraic manipulation, we have that:

$$\hat{\mathbf{s}} = \mathbf{H}_{s_1} \mathbf{s} + \mathbf{H}_{s_2} \mathbf{s}^* \quad (37)$$

where

$$\mathbf{H}_{s_1} = \frac{1}{2} \begin{bmatrix} \mathbf{h}_{x_1 x_1}^T + \mathbf{h}_{x_Q x_Q}^T + j(\mathbf{h}_{x_Q x_1}^T - \mathbf{h}_{x_1 x_Q}^T) \\ \mathbf{h}_{y_1 x_1}^T + \mathbf{h}_{y_Q x_Q}^T + j(\mathbf{h}_{y_Q x_1}^T - \mathbf{h}_{y_1 x_Q}^T) \\ \mathbf{h}_{x_1 y_1}^T + \mathbf{h}_{x_Q y_Q}^T + j(\mathbf{h}_{x_Q y_1}^T - \mathbf{h}_{x_1 y_Q}^T) \\ \mathbf{h}_{y_1 y_1}^T + \mathbf{h}_{y_Q y_Q}^T + j(\mathbf{h}_{y_Q y_1}^T - \mathbf{h}_{y_1 y_Q}^T) \end{bmatrix} \quad (38)$$

$$\mathbf{H}_{s_2} = \frac{1}{2} \begin{bmatrix} \mathbf{h}_{x_1 x_1}^T - \mathbf{h}_{x_Q x_Q}^T + j(\mathbf{h}_{x_Q x_1}^T + \mathbf{h}_{x_1 x_Q}^T) \\ \mathbf{h}_{y_1 x_1}^T - \mathbf{h}_{y_Q x_Q}^T + j(\mathbf{h}_{y_Q x_1}^T + \mathbf{h}_{y_1 x_Q}^T) \\ \mathbf{h}_{x_1 y_1}^T - \mathbf{h}_{x_Q y_Q}^T + j(\mathbf{h}_{x_Q y_1}^T + \mathbf{h}_{x_1 y_Q}^T) \\ \mathbf{h}_{y_1 y_1}^T - \mathbf{h}_{y_Q y_Q}^T + j(\mathbf{h}_{y_Q y_1}^T + \mathbf{h}_{y_1 y_Q}^T) \end{bmatrix} \quad (39)$$

which leads to the widely linear MIMO complex-valued equalizer structure

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{xx}^H & \mathbf{h}_{xy}^H \\ \mathbf{h}_{yx}^H & \mathbf{h}_{yy}^H \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + \begin{bmatrix} \mathbf{h}_{xx}^{H*} & \mathbf{h}_{xy}^{H*} \\ \mathbf{h}_{yx}^{H*} & \mathbf{h}_{yy}^{H*} \end{bmatrix} \begin{bmatrix} \mathbf{x}^* \\ \mathbf{y}^* \end{bmatrix} \quad (40)$$

The corresponding block diagram for the equalizer is depicted in Fig. 1.

Observe that Eq. (40) reduces to Eq. (34) (i.e. $\mathbf{H}_{s_2} = 0$) if there is no mixing between IQ components in the received signal. One advantage of the structure in Eq. (40) over the one in Eq. (32) is that, since blind equalization algorithms usually rely on the two-dimensional statistical properties of complex-valued signals, those algorithms are easier to derive for a complex-valued equalizer.

B. Filter Coefficients Update Algorithm

Taking as start point the the complex least mean square (CLMS) algorithm, the widely linear CLMS can be derived (WL-CLMS) [14]. The adaptation rules used to train a the WL-CLMS are similar to the ones obtained for the CLMS, as shown in [15] and [14].

However, in order to make the dynamic update of the FIR coefficients independent of the carrier phase, the update rules for the widely linear MIMO equalizer used in this paper were implemented based on the algorithm described in [15], modified to use a decision directed least radius distance (DD-LRD) error criteria [16]. This error criteria can be easily changed to implement a radius directed equalizer (RDE) mode, in order to perform unsupervised training of the equalizer's coefficients. The filters of \mathbf{H}_{s_1} and \mathbf{H}_{s_2} are updated according to Eq. (41) and (42), respectively:

$$\begin{aligned} \mathbf{h}_{xx}(k+1) &= \mathbf{h}_{xx}(k) + \mu e_x(k) \mathbf{x}^*(k) \\ \mathbf{h}_{xy}(k+1) &= \mathbf{h}_{xy}(k) + \mu e_x(k) \mathbf{y}^*(k) \\ \mathbf{h}_{yx}(k+1) &= \mathbf{h}_{yx}(k) + \mu e_y(k) \mathbf{x}^*(k) \\ \mathbf{h}_{yy}(k+1) &= \mathbf{h}_{yy}(k) + \mu e_y(k) \mathbf{y}^*(k) \end{aligned} \quad (41)$$

$$\begin{aligned} \mathbf{h}_{xx}^*(k+1) &= \mathbf{h}_{xx}^*(k) + \mu e_x(k) \mathbf{x}(k) \\ \mathbf{h}_{xy}^*(k+1) &= \mathbf{h}_{xy}^*(k) + \mu e_x(k) \mathbf{y}(k) \\ \mathbf{h}_{yx}^*(k+1) &= \mathbf{h}_{yx}^*(k) + \mu e_y(k) \mathbf{x}(k) \\ \mathbf{h}_{yy}^*(k+1) &= \mathbf{h}_{yy}^*(k) + \mu e_y(k) \mathbf{y}(k) \end{aligned} \quad (42)$$

where $0 < \mu < 1$ is the adaptation step, and $\{e_x(k), e_y(k)\}$ correspond to the error criteria, defined by

$$\begin{aligned} e_x(k) &= \hat{x}(k)(|x_d(k)|^2 - |\hat{x}(k)|^2) \\ e_y(k) &= \hat{y}(k)(|y_d(k)|^2 - |\hat{y}(k)|^2) \end{aligned} \quad (43)$$

where $x_d(k)$ and $y_d(k)$ are training symbols, or

$$\begin{aligned} e_x(k) &= \hat{x}(k)(|R_x(k)|^2 - |\hat{x}(k)|^2) \\ e_y(k) &= \hat{y}(k)(|R_y(k)|^2 - |\hat{y}(k)|^2) \end{aligned} \quad (44)$$

where $R_x(k)$ and $R_y(k)$ are the decided radius of the RDE algorithm. Both error criteria are independent of frequency offset.

V. SIMULATION SETUP AND NUMERICAL RESULTS

We numerically investigate the performance of the widely linear adaptive equalizer for dual polarization (DP) transmission using M -QAM modulation formats, with $M = 16, 64$ and 256. The numerical simulation model is illustrated in Fig. 2. For all simulated cases, we assumed a symbol rate (R_s) of 32 GBd. In the constellation mapping block, the sequences of M -QAM symbols were generated with a fixed length of 250000 symbols per polarization, mapped from decorrelated pieces of pseudo-random bit sequences (PRBS, $2^{23} - 1$). The modulated data was upsampled to 16 samples per symbol, and pulse shaped using root raised cosine (RRC) filters with rolloff factor of 0.01. For all tested cases, a fixed frequency offset of 100 MHz was assumed between the local oscillators at transmitter and receiver. No phase noise was assumed.

The simplified linear optical channel model used accounted for static linear polarization rotation, CD loading, and amplified spontaneous emission (ASE) from Erbium doped fiber amplifiers (EDFAs), modeled as additive white Gaussian noise (AWGN) loading stage. At the receiver front-end model, the frequency offset is added to the signal, which is filtered by an ideal rectangular low pass filter with a bandwidth of $1.1R_s$. This stage is followed by an ideal dual polarization coherent receiver, and a block to simulate IQ imbalance and IQ skew on the received polarizations. After the receiver front-end, a standard set of DSP algorithms for data demodulation: frequency domain CD compensation, decimation, adaptive equalization, ideal carrier recovery (i.e., 100 MHz frequency offset compensation) and digital demodulation using hard symbol decisions with minimum Euclidean distance.

The adaptive equalizers were implemented in the $T_s/2$ -fractional spaced mode, where $T_s = 1/R_s$, therefore requiring the signal to be sampled at $2R_s$.

For all results that follows, the label “complex” refers to the 2×2 complex-valued (strictly) linear equalizer structure defined in Eq. (34) and the label “WL complex” refers to the 2×2 complex-valued widely linear structure defined in Eq. (40). The chosen figure of merit for the receiver's performance comparisons is the Q^2 -factor, in dB, which is calculated from the bit error rate (BER) according to $Q_{dB}^2 = 20 \log_{10}[\sqrt{2} \text{erfcinv}(2BER)]$. Distinct optical signal-to-noise ratio (OSNR) values were chosen for each tested modulation format in order to maintain the minimum BER values around 10^{-4} to 10^{-3} at the receiver. The OSNR is defined assuming a bandwidth of 12.5 GHz.

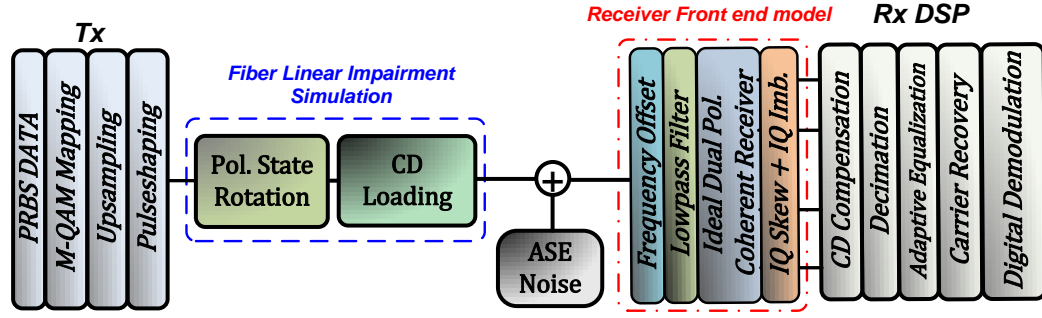
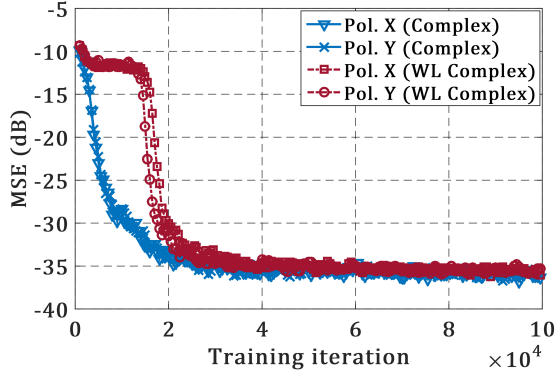


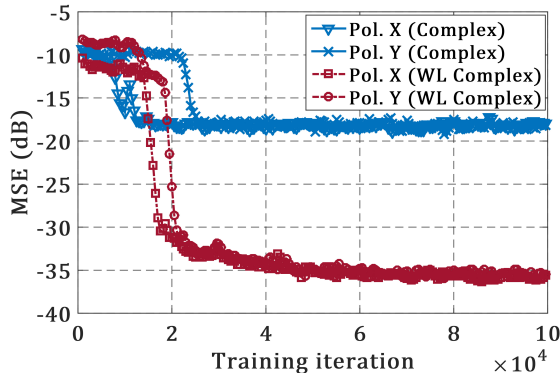
Fig. 2 Numerical simulation setup of a simplified optical channel, with only fiber linear propagation effects and additive white Gaussian noise included.

A. MSE Convergence

Figure 3 shows the convergence of the equalizers in terms of mean square error (MSE) of the training iteration. Both equalizers were configured with 25 taps time response. For this particular example, each point correspond the mean square error calculated with a moving average window of 1000 iterations. These results refer to the polarization demultiplexing of a 32 GBd DP-256QAM signal in the “back-to-back” configuration with 37 dB of OSNR at the receiver’s input.



(a) MSE evolution in absence of IQ-mixing effects.



(b) MSE evolution for both polarizations subject to: $\epsilon_{rx} = 0.05$ (IQ amplitude imbalance), $\phi_{rx} = 4^\circ$ (IQ phase imbalance), and $\tau = 0.25T_s$ (IQ skew).

Fig. 3 Example of MSE evolution during the equalizers’ training iterations.

Figure 3 (a) corresponds to the case where no IQ imbalance or IQ skew is present at the receiver front-end. The MSE of both equalizers converges to a minimum value of -35 dB. Figure 3 (b) corresponds to the case where the signals of both

polarizations are subject to IQ imbalance and IQ skew. For this case, it can be noticed that the MSE of the 2×2 complex-valued linear equalizer shows an offset of more than 15 dB, compared with the curve in Fig. 3 (a). However, the same floor of MSE of -35 dB is reached with 2×2 complex-valued widely linear equalizer, regardless the presence or not of IQ-mixing.

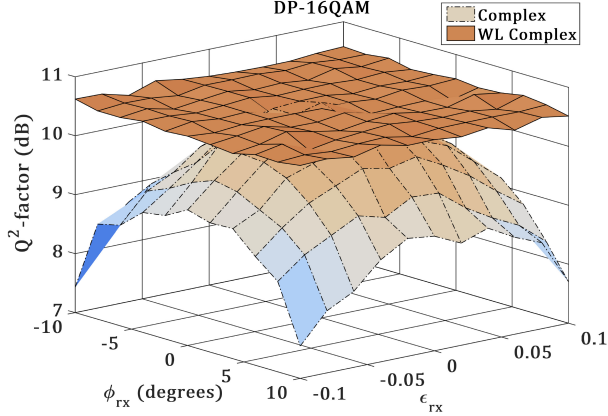
The training period correspond to sequence of consecutive tap updates, with each iteration using one training symbol. In Fig. 3 it is shown the convergence of the equalizer for a training sequence (block) of 10×10^4 symbols. It can be noticed that the convergence is achieved in between 20×10^3 and 40×10^3 iterations, which provides an estimate of the required size of the training sequence. Additionally, it is also clear that both equalizers converge in the MSE with more or less the same number of iterations, which indicates that the complex-valued WL implementation did not increase the requirements of training length compared to the conventional strictly linear complex-valued structure.

B. Performance with IQ Imbalance and Skew in Back-to-back

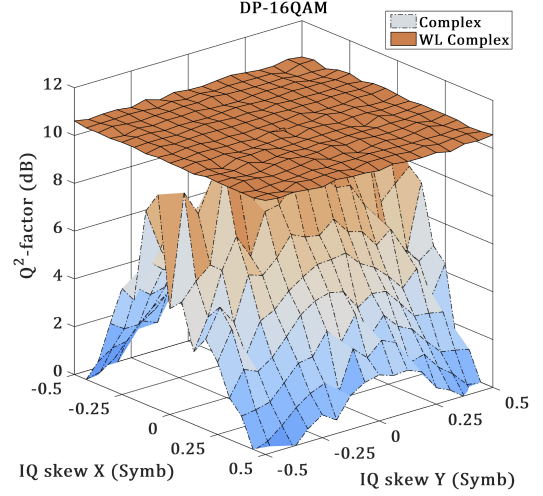
In Fig. 4 the performance surfaces of the receiver for the two complex-valued adaptive equalizers tested for different combinations of IQ amplitude imbalance (ϵ_{rx}), IQ phase imbalance (ϕ_{rx}), and IQ skew. The performance is evaluated in the stationary regime, i. e. after convergence in the MSE sense during the training period. Both equalizers were configured to use 25 taps. As it can be seen, in presence of any of the impairments, the performance of the 2×2 complex-valued equalizer is degraded. The penalty in Q^2 -factor, as one would expect, also increases with the order of the modulation format.

On the other hand, for all tested cases, the optimal performance is maintained using 2×2 complex-valued WL equalizer. Here, we assume as “optimal” the receiver performance after 2×2 complex-valued linear equalization in absence of any IQ-mixing at the receiver front-end.

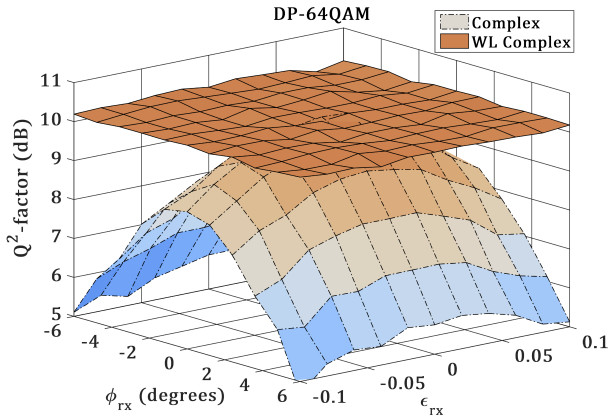
The spikes appearing in the performance surfaces of the 2×2 complex-valued linear equalizer are attributed to convergence to the same polarization. Since this structure cannot handle IQ-mixing, those extreme cases where one polarization is perfectly received while the other is severely impaired by IQ-skew may have an increasing probability of convergence to the same polarization. However, an additional study would be necessary to precisely characterize the cause and frequency of such events.



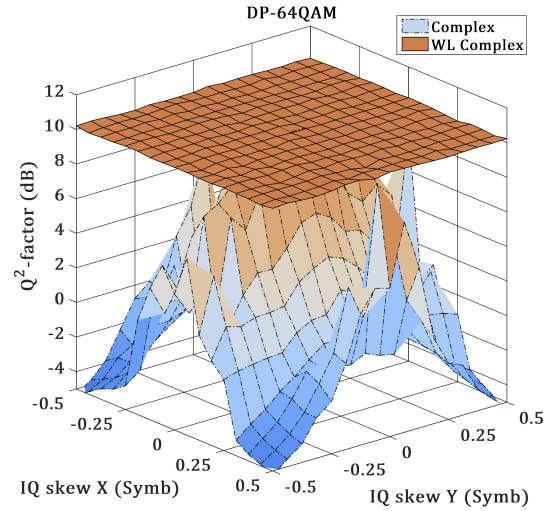
(a) Q^2 -factor vs normalized IQ amplitude imbalance (ϵ_{rx}) and IQ phase imbalance (ϕ_{rx}). Same values applied in both polarizations at the receiver front-end, for 32 GBd DP-16QAM with $OSNR_{RX} = 22$ dB.



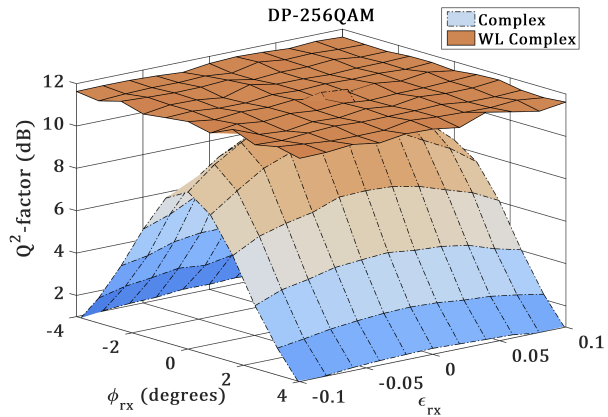
(b) Q^2 -factor vs IQ skew DP-16QAM, $OSNR_{RX} = 22$ dB.



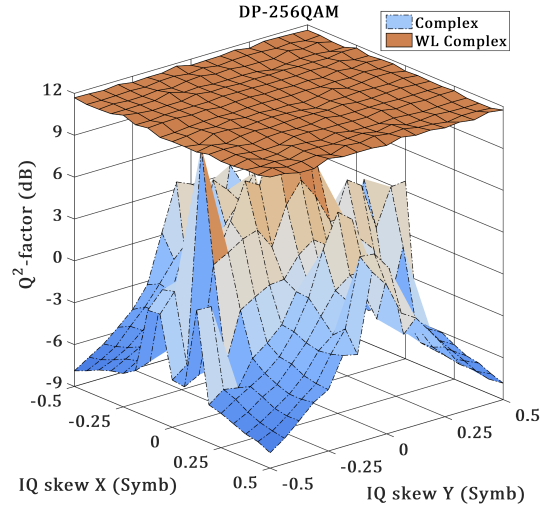
(c) Q^2 -factor vs normalized IQ amplitude imbalance (ϵ_{rx}) and IQ phase imbalance (ϕ_{rx}). Same values applied in both polarizations at the receiver front-end, for 32 GBd DP-64QAM with $OSNR_{RX} = 28$ dB.



(d) Q^2 -factor vs IQ skew DP-64QAM, $OSNR_{RX} = 28$ dB.



(e) Q^2 -factor vs normalized IQ amplitude imbalance (ϵ_{rx}) and IQ phase imbalance (ϕ_{rx}). Same values applied in both polarizations at the receiver front-end, for 32 GBd DP-256QAM with $OSNR_{RX} = 37$ dB.



(f) Q^2 -factor vs IQ skew DP-256QAM, $OSNR_{RX} = 37$ dB.

Fig. 4 Back-to-back performance after 2×2 complex-valued equalization and after 2×2 widely linear complex-valued equalization for a dual polarization coherent receiver subject to IQ imbalance or IQ skew.

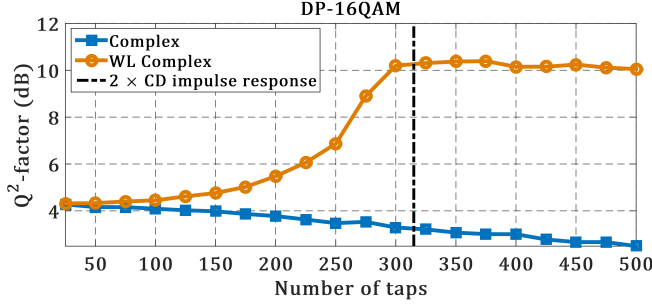
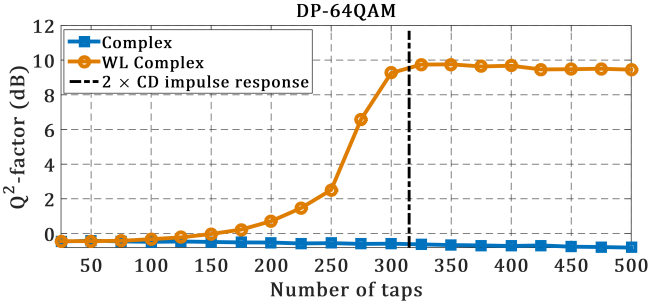
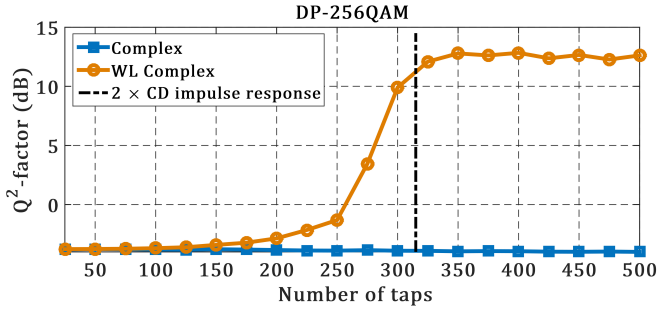
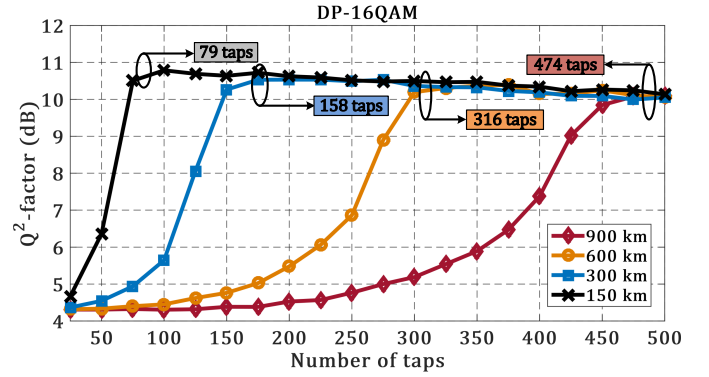
(a) Q^2 -factor vs number of equalizer taps DP-16QAM, $\text{OSNR}_{RX} = 22$ dB.(b) Q^2 -factor vs number of equalizer taps DP-64QAM, $\text{OSNR}_{RX} = 28$ dB.(c) Q^2 -factor vs number of equalizer taps DP-256QAM, $\text{OSNR}_{RX} = 37$ dB.

Fig. 5 Receiver performance comparison after zero-forcing CD compensation of 9600 ps/nm, equivalent to 600 km of transmission over a single mode fiber with dispersion parameter 16 ps/nm/km, for different modulation formats.

C. Performance with IQ Imbalance and Skew after Zero-forcing CD Compensation

We then compare the receiver performance when the signal is directed to the adaptive equalizer after zero-forcing equalization of a large amount of CD. For all simulated cases, the assumption is that the receiver front-end adds $\epsilon_{rx} = 0.05$ of IQ amplitude imbalance, $\phi_{rx} = 4^\circ$ of IQ phase imbalance, and $\tau = 0.25T_s$ of IQ skew, for both received polarizations. The parameters of the CD loading stage were configured to provide a total of 9600 ps/nm of accumulated CD, which is equivalent to a transmission over 600 km of single mode fiber with a dispersion parameter of 16 ps/nm/km at 1550 nm.

Using the analytic expressions in [17], we can calculate that a 32 Gb/s signal, sampled at 64 GS/s, would require from a $T/2$ -fractional spaced equalizer (FSE) a minimum of 158 taps to compensate for such amount of CD. The value corresponding to twice of this lower bound, 316 taps, is

Fig. 6 Receiver performance of DP-16QAM, $\text{OSNR}_{RX} = 22$ dB, using the adaptive WL equalizer after zero-forcing CD compensation. Each curve refer to a fixed transmission distance (dispersion parameter 16 ps/nm/km) and indicate the Q^2 -factor as a function of the equalizer number of taps. The values pointed by the arrows indicate the corresponding number of taps to cover twice of CD response duration for each respective distance.

indicated by the vertical dashed lines in the subfigures of Fig. 5.

The results shown in Fig. 5 demonstrate that the complex-valued 2×2 equalizer is not able to recover the signal with any number of taps configuration. On the other hand, the performance of the 2×2 complex-valued WL equalizer clearly depends on its number of taps. In particular, when the number of taps exceed 316 (i.e., twice the CD impulse response), the performance of the receiver is identical to the one obtained with no CD loading. Therefore, these results are in agreement with the conclusions drawn from the analysis presented in the section III. Finally, Fig. 6 shows that the same behaviour of the 2×2 complex-valued WL equalizer observed in Fig. 5 is reproduced for different transmission distances.

D. Additional Comments

As the results have shown, a widely linear equalization structure can be used to compensate practically all reversible combinations of IQ imbalance and IQ skew present at the receiver. However, such capability doubles the complexity of the standard channel equalization. Therefore, considering how the requirements increase with the amount accumulated CD in the transmission link we can point out that, for long haul systems, the overhead of complexity may be too large for practical implementation. Nevertheless, the models derived in section III can be used to investigate the impact of receiver front-end imperfections on the performance of long-haul transceivers, which is useful information for equipment calibration and for the development of algorithms to estimate and compensate IQ-mixing effects before the DSP equalization blocks. On the other hand, for metro networks and short range transceivers, where the amount of accumulated CD can be reasonably small, such equalizer structure tend to increase its cost effectiveness, since the use of widely linear equalization may relax the requirements on the subsystems components used in coherent receivers, for example. For the last, the study of the symmetries of the complex-valued WL models may provide insights to optimize its DSP implementations.

VI. CONCLUSION

In this paper we have shown that using augmented complex analysis it is possible to derive useful complex-valued models of IQ imbalance and IQ skew in optical coherent receivers. We have shown how this models naturally connect the compensation of such impairments and the general equalization problem of the linear channel response. It is shown that the attempt to use a conventional zero-forcing CD equalizer under the presence of IQ-mixing receiver front-end imperfections results in a two fold increase in the complexity required by the adaptive equalization algorithm, with respect to the static CD equalizer, independently of the implementation domain (real-valued space or complex-valued space). Based on the augmented complex analysis, we also have proposed and extensively validated by numerical simulations, for different M -QAM modulation formats, a new widely linear complex-valued adaptive equalizer structure for a polarization multiplexed receiver. The proposed structure is able to maintain optimal performance, in a MMSE sense, under the presence of IQ imbalance and IQ skew at the receiver front-end. Results from extensive numerical simulations have shown that the performance of the proposed equalizer agrees with the analytical models first derived.

ACKNOWLEDGMENT

The authors would like to thank the financial support from the VILLUM FOUNDATION, Søborg, Denmark.

REFERENCES

- [1] Seb J. Savory, "Digital Coherent Optical Receivers: Algorithms and Subsystems," in *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164 - 1179, 2010.
- [2] R. W. Lucky, "Automatic equalization for digital communication," in *The Bell System Technical Journal*, vol. 44, no. 4, pp. 547 - 588, 1965.
- [3] Paskov, M.; Lavery, D.; Savory, S.J., "Blind Equalization of Receiver In-Phase/Quadrature Skew in the Presence of Nyquist Filtering," in *IEEE Photon. Technol. Lett.*, vol. 25, no. 24, pp. 2446 - 2449, 2013.
- [4] M. S. Faruk and K. Kikuchi, "Compensation for in-phase/quadrature imbalance in coherent-receiver front-end for optical quadrature amplitude modulation," *IEEE Photonics J.*, vol. 5, no. 2, 2013.
- [5] R. Rios-Müller, J. Renaudier, G. Charlet, "Blind Receiver Skew Compensation for Long-Haul Non-Dispersion Managed Systems," *Eur. Conf. Opt. Commun.*, pp. 4 - 6, 2014.
- [6] R. Rios-Müller, J. Renaudier, and G. Charlet, "Blind Receiver Skew Compensation and Estimation for Long-Haul Non-Dispersion Managed Systems Using Adaptive Equalizer," in *J. Light. Technol.*, vol. 33, no. 7, pp. 13151318, 2015.
- [7] Karlsson, M., "Four-dimensional Rotations in Coherent Optical Communications," in *J. Light. Technol.*, vol. 32, no. 6, pp. 1246-1257, 2014.
- [8] P. Rykaczewski, M. Valkama, and M. Renfors, "On the Connection of I/Q Imbalance and Channel Equalization in Direct-Conversion Transceivers," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1630 - 1636, 2008.
- [9] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-Valued Signal Processing: The Proper Way to Deal With Improperity," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101 - 5125, 2011.
- [10] D. Mittera, L. Paura and F. Sterle, "Widely linear MMSE equalizer for MIMO linear time-dispersive channel," in *Proc. of the 3rd International Symposium on Image and Signal Processing and Analysis*, 2003.
- [11] D. Darsena, G. Gelli, L. Paura and F. Verde, "Widely linear equalization and blind channel identification for interference-contaminated multicarrier systems," in *IEEE Trans. Signal Process.*, Vol. 53, No. 3, pp. 1163 - 1177, 2005.
- [12] D. Mittera, L. Paura and F. Sterle, "Widely linear decision-feedback equalizer for time-dispersive linear MIMO channels," in *IEEE Trans. Signal Process.*, Vol. 53, No. 7, pp. 2525 - 2536, 2005.
- [13] B. S. Chang, C. Rocha, D. Ruyet and D. Roviras, "Widely linear MMSE precoding and equalization techniques for SC-FDE systems," *EURASIP Journal on Advances in Signal Processing*, 2014.
- [14] Mandic, Danilo and Goh, Vanessa Su Lee, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, Wiley Publishing, 2009.
- [15] Soroush Javidi and Maciej Pedzisz and Su Lee Goh and Danilo P. Mandic, "The Augmented Complex Least Mean Square Algorithm With Application To Adaptive Prediction Problems" in *Proc. 1st IARP Workshop on Cognitive Information Processing*, 2008.
- [16] X. Xu, B. Chtelain, and D. V. Plant, "Decision Directed Least Radius Distance Algorithm for Blind Equalization in a Dual-polarization 16-QAM System," in *Proc. Optical Fiber Communication Conference (OFC 2012)*, paper OM2H.5, 2012.
- [17] Ip, E., Kahn, J.M., "Digital Equalization of Chromatic Dispersion and Polarization Mode Dispersion," in *J. Light. Technol.*, vol. 25, no. 8, pp. 2033 - 2043, 2007.

Edson Porto da Silva was born in Pocinhos, Brazil, on August 9, 1988. He received the B.S. degree in Electrical Engineering from Federal University of Campina Grande (UFCG) in Campina Grande, Paraíba, Brazil, in 2011, and the M.S. degree in Electrical Engineering from State University of Campinas (UNICAMP), Campinas, São Paulo, Brazil, in 2013. From July 2011 to November 2013, he was with the Center for Research and Development in Telecommunications (CPqD), Campinas, working as researcher in coherent optical communications and focused on experimental demonstrations of 400 Gb/s WDM systems. He is currently pursuing his Ph.D. degree at the Photonics Engineering Department (Fotonik), Technical University of Denmark (DTU), Kgs. Lyngby, Denmark. His research interests include digital signal processing for coherent optical receivers, mitigation of fiber non-linear impairments, high speed digital optical transmission, and machine learning techniques

Darko Zibar was born in Belgrade, Serbia, on September 9, 1978. He received the M.Sc. degree in telecommunication and the Ph.D. degree in optical communications from the Technical University of Denmark, Lyngby, Denmark, in 2004 and 2007, respectively. He was a Visiting Researcher with the Optoelectronic Research Group, University of California, Santa Barbara, CA, USA, from January 2006 to August 2006, and in January 2008, where he worked on coherent receivers for phase-modulated analog optical links. From February 2009 to July 2009, he was a Visiting Researcher with Nokia-Siemens Networks, where he worked on 112 Gb/s polarization multiplexed systems. He is currently an Associate Professor at DTU Fotonik, Technical University of Denmark. His research interests include coherent optical communication, statistical signal processing, and Bayesian inference. He received the Best Student Paper Award at the IEEE Microwave Photonics Conference 2006, the Villum Kann Rasmussen Postdoctoral Research Grant in 2007, and the Villum Foundation Young Investigator Program in 2012.